WHAT IS CLAIMED IS:

1. A text classification apparatus comprising:

a text input device for receiving an entered text;

a storage device for storing a concept thesaurus file for use in classifying an entered text to be classified, a cross lingual word sense-based knowledge file corresponding to a plurality of languages including a first and a second language, and a word-based classification knowledge file;

a processing unit for executing a classification of the entered text to be classified to assign a category to the text; and

an output device for outputting the classification result, wherein:

said text input device receives an entered text to be classified in the first language,

said processing unit is configured to:

extract a word from said first text to be classified;

convert the extracted a word into a word sense using said concept thesaurus file;

compare the word sense resulting from the conversion with information on each category included in said cross lingual word sense-based classification knowledge file to calculate a first score for each category;

compare the extracted word with word

classification information included in said word-based classification knowledge file to calculate a second score for each category; and

integrate said first and second scores for each category to determine a category for the first text to be classified in the first language for assigning a category to the first text, and

said word-based classification knowledge file is generated by learning a word-based classification knowledge using words included in a labeled text in the first language.

2. A text classification apparatus according to claim 1, wherein:

said text to be classified which has been assigned the category by said text classification apparatus is used for learning the word-based classification knowledge as a labeled text in the first language used in the generation of said word-based classification knowledge file.

3. A text classification apparatus according to claim 1, wherein:

said cross lingual word sense-based classification knowledge file is generated by extracting a word included in a labeled text in the second language, converting the extracted word into a word sense using said concept thesaurus file, and learning the word sense-based classification knowledge using said word sense and category information on the

labeled text in the second language.

4. A text classification apparatus according to claim 1, further comprising a display device, and a user input device for accepting an entry from a user, wherein:

said display device displays a word sense used in the assignment of a category, and displays together word sense candidates other than the word sense used in the assignment of a category when a plurality of word sense candidates exist for a single word extracted from the text to be classified, and

said text classification apparatus receives information entered for selecting one of the plurality of word sense candidates through said user input device, changes a word sense which is compared with information on each category included in said cross lingual word sense-based classification knowledge file based on the entered selection information to assign a category to the text to be classified.

5. A text classification apparatus according to claim 1, further comprising a display device, and a user input device for accepting an entry from a user, wherein:

said processing unit compares a first word sense included in said cross lingual word sense-based classification knowledge file or said word-based classification knowledge file for the category assigned to the text to be classified with a second word sense

resulting from a conversion of a word included in the text to be classified by said concept thesaurus to extract contradictory word senses, and

said display device displays the first and second word senses for the extracted word, and receives information entered for selecting one of said first and second word senses through said user input device; and

said text classification apparatus changes a word sense corresponding to each category included in said cross lingual word sense-based classification knowledge file or said word-based classification knowledge file based on the entered selection information.

6. A text classification apparatus comprising:

a text input device for receiving an entered text;

a storage device for storing a concept thesaurus file for use in classifying an entered text to be classified, a cross lingual word sense knowledge file corresponding to a plurality of languages including a first and a second language, and a word-based classification knowledge file;

a processing unit for executing a classification of the entered text to be classified to assign a category to the text; and

an output device for outputting the classification result, wherein:

said text input device receives an entered

text to be classified in the first language,

said processing unit is configured to:

extract a word from said first text to be classified;

convert the extracted word into a word sense using said concept thesaurus file;

compare the word sense resulting from the conversion with information on each category included in said cross lingual word sense-based classification knowledge file to calculate a first score for each category;

compare the extracted word with word-based classification information included in said word-based classification knowledge file to calculate a second score for each category; and

integrate said first and second scores for each category to determine a category for the first text to be classified in the first language for assigning a category to the first text, and

said word-based classification knowledge file is generated by extracting information indicative of a relation between a plurality of words from a labeled text in the first language, and extracting a word-based classification knowledge of each category using the extracted information on the relation between words, and word classification information on each category included in said cross lingual word sense-based classification knowledge file.

7. A text classification apparatus according to claim 6, wherein:

said cross lingual word sense-based classification knowledge file is generated by extracting a word included in a text in the second language with category information, converting the extracted word into a word sense using said concept thesaurus, and learning a word sense-based classification knowledge using the word sense and category information of the labeled text in the second language.

8. A text classification apparatus according to claim 6, further comprising a display device, and a user input device for accepting an entry from a user, wherein:

said display device displays a word sense used in the assignment of a category, and displays together word sense candidates other than the word sense used in the assignment of a category when a plurality of word sense candidates exist for a single word extracted from the text to be classified; and

receives information entered for selecting one of the plurality of word sense candidates through said user input device, changes a word sense which is compared with information on each category included in said cross lingual word sense-based classification knowledge file based on the entered selection information to assign a category to the text to be

classified.

9. A text classification apparatus according to claim 6, further comprising a display device, and a user input device for accepting an entry from a user, wherein:

said processing unit compares a first word sense included in said cross lingual word sense-based classification knowledge file or said word-based classification knowledge file for the category assigned to the text to be classified with a second word sense resulting from a conversion of a word included in the text to be classified by said concept thesaurus to extract contradictory word senses, and

said display device displays the first and second word senses for the extracted word, and receives information entered for selecting one of said first and second word senses through said user input device; and

said text classification apparatus changes a word sense corresponding to each category included in said cross lingual word sense-based classification knowledge file or said word-based classification knowledge file based on the entered selection information.

10. A method of assigning a category to a text to be classified, said text being entered into a text classification apparatus having a storage device for storing a concept thesaurus file for use in classifying an entered text to be classified, a cross lingual word

sense knowledge file corresponding to a plurality of languages including a first and a second language, and a word-based classification knowledge file, and a processing unit for executing a classification of the entered text to be classified to assign a category to the text, said method comprising the steps of:

receiving a text in the first language entered for classification;

extracting a word from the text to be classified in the first language;

converting the extracted word into a word sense using said concept thesaurus file;

comparing the word sense resulting from the conversion with information on each category included in said cross lingual word sense-based classification knowledge file to calculate a first score for each category;

comparing the extracted word with word-based classification information included in said word-based classification knowledge file to calculate a second score for each category; and

integrating said first and second scores for each category to determine a category for the text to be classified in the first language for assigning a category to the text to be classified, and

said word-based classification knowledge file being generated by learning a word-based classification knowledge using words included in the labeled text in

the first language.

11. A text classification method according to claim 10, wherein:

said text to be classified which has been assigned the category by said text classification apparatus is used for learning the word-based classification knowledge as the labeled text in the first language used in the production of said word-based classification knowledge file.

12. A text classification method according to claim 10, wherein:

said cross lingual word sense-based classification knowledge file is generated by extracting a word included in a labeled text in the second language, converting the extracted word into a word sense using said concept thesaurus file, and learning the word sense-based classification knowledge using said word sense and category information on the labeled text in the second language.

13. A text classification method according to claim 10, wherein said text classification apparatus further has a display device, and a user input device for accepting an entry from a user, said method further comprising the steps of:

displaying a word sense used in the assignment of a category on said display device;

displaying together word sense candidates other than the word sense used in the assignment of a

category when a plurality of word sense candidates exist for a single word extracted from the text to be classified;

receiving information entered for selecting one of the plurality of word sense candidates through said user input device; and

changing a word sense which is compared with information on each category included in said cross lingual word sense-based classification knowledge file based on the entered selection information to assign a category.

14. A text classification method according to claim 10, wherein said text classification apparatus further has a display device, and a user input device for accepting an entry from a user, said method further comprising the steps of:

comparing a first word sense included in said cross lingual word sense-based classification knowledge file or said word-based classification knowledge file for the category assigned to the text to be classified with a second word sense resulting from a conversion of a word included in the text to be classified by said concept thesaurus to extract contradictory word senses;

displaying the first and second word senses for the extracted word, and receiving information entered for selecting one of said first and second word senses through said user input device; and

changing a word sense corresponding to each

category included in said cross lingual word sense-based classification knowledge file or said word-based classification knowledge file based on the entered selection information.